Iranian Society of Acoustics and Vibration
The 12th ISAV2022 International Conference on Acoustics and Vibration
14-15 Dec 2022 Tehran - Iran
Amirkabir University of Technology

Detection of Blade Crack in a Rotor System Using MLP-Based Automatic Feature Selector

Emadaldin Sh Khoram-Nejad[a], Abdolreza Ohadi[b*]

[a] *PhD student, Acoustics Research Laboratory, Mechanical Engineering Department, Amirkabir University of Technology (Tehran Polytechnic), Tehran, Iran.*
[b] *Professor, Acoustics Research Laboratory, Mechanical Engineering Department, Amirkabir University of Technology (Tehran Polytechnic), Tehran, Iran.*
** Corresponding author e-mail: a_r_ohadi@aut.ac.ir*

Abstract

This paper aimed to detect the crack location in a shaft-disk-cracked blade (SDCB) system. An open crack with three different depths and three distances from the blade root is considered. A finite element (FE) simulation of the SDCB system rotating at 3000 rpm was done in the MSC Adams as commercial software, simulating the systems' dynamics. The results are compared with those of another FE software and the assumed mode method (AMM) to validate the modeling procedure. Feature extraction and selection always have crucial roles in classification problems. Therefore, this paper tries to show the benefit of automatic feature selection compared to manual feature selection to detect the crack location. For this purpose, the performance of an ensemble classifier using manual feature selection is compared to a deep multi-layer perceptron (MLP)-based classifier, which selects the best features automatically and determines the crack location. The ensemble classifier includes the support vector machine (SVM), k-nearest neighbor (k-NN), and decision tree classifiers. Then, the sensitivity of the proposed crack detection models to different measuring points, one at the bearing location and another near the cracked blade tip, is studied. The best location for sensors to detect cracks more accurately is determined during this comparison. The results indicate that the best accuracy of crack detection is achieved when using the MLP-based classifier and the signals of the displacement of the cracked blade tip. This accuracy is 97.14%, and the existence of a crack is detected without an error (100% accuracy).

Keywords: Crack detection; Feature selection; Assumed mode method; Multi-layer perceptron; Ensemble classifier

1. Introduction

Nowadays, rotating machines can be found everywhere, in large power plants or small workshops. According to the applications, these machines rotate at speeds ranging from hundreds to ten thousand

rpm. From a safety point of view, there is no fully healthy machine in the industry due to manufacturing procedures and microstructural faults. Also, considering its environment and working condition, some faults appear in the machine. One of the common faults in the shaft-disk-blade systems, such as compressors, is the existence and growing of cracks on the blade.

Fault detection of the rotating machines is an exciting subject in mechanical engineering which is almost always an industrial requirement. In contrast to faults like misalignment or unbalance, which significantly affect the signal's dynamic behavior and frequency response function (FRF), crack detection is challenging due to its weak characteristics in vibrational signals. It is a well-known concept because the crack fault is always the reason for the sudden failure of the systems.

Artificial intelligence has revolutionized many fields of study, and the fault diagnosis of rotating machines has no exception [1]. Machine learning can be categorized as a data-based solution. It means that there is no need to have a closed-form formula (as in analytical methods) or discretize the whole domain of the systems and solve the problem for sub-domains (as in numerical methods). Machine learning methods consider a model that is supposed to be trained with a training dataset and recognize the pattern behind the data. Then this model is tested with a new related dataset, and the training model's ability on the new dataset shows the training model's effectiveness. The machine learning method uses labeled datasets for the fault diagnosis procedure and tries to classify the data.

However, there is some limitation in the input of these data-based methods. Almost all machine learning classifiers, such as support vector machines (SVM), require more valuable and compact inputs and try to learn all of the specific features from the input datasets. Suppose the input signal is raw, based on the fact that the raw signal contains significant amounts of noise and information unrelated to a specific fault, these methods will not perform well and cannot identify the fault correctly. The solution is to use feature extraction and selection procedure techniques. Using some popular statistic or probabilistic properties of signal in time, frequency, or time-frequency domain such as kurtosis or standard deviation can help the classifiers detect faults more precisely.

Researches have been done to diagnose rotating machinery faults using different methods. Bearing and gearbox faults [2,3] are among the most famous fault diagnosis case studies. Some research has been done on rotor faults diagnoses such as unbalance [4,5], rub-impact [6,7], and misalignment [8]. Among all the literature in the field of fault diagnosis of rotating machines, few researchers, such as [9–12], have worked on blade crack detection. Therefore, network-based crack detection methods need more attention to provide more accurate models to precisely detect the crack and its location.

For detecting a crack in a system using data-based models, such as machine learning or deep learning models, it is necessary to have a dataset from which the model can extract the characteristics of the crack. Therefore, the location and the type of the sensors used in the data acquisition step of crack detection is an important decision. Although the traditional way to capture the signals is to use accelerometers on the bearings or strain gauges on the blades, another way is to use blade tip timing methods which have been widely used in recent years [13,14].

This paper proposes a deep MLP-based classifier with manual and automatic feature selection to detect the crack's existence and location in a shaft-disk-cracked-blade (SDCB) system. Due to the rare research on blade crack detection using neural networks and especially detecting the crack location in each blade, this paper proposes a novel approach to select the beneficial features automatically. Also, its performance is compared with an ensemble classifier to show the superiority of the proposed classifier. Another investigation of this paper is to find and suggest the best location to capture the more beneficial signals for crack detection purposes, and this suggestion is practically possible.

2. Finite element simulation of the SDCB system

The SDCB system is modeled in the MSC Adams as a well-known dynamic mechanism modeling commercial software (Figure 1).

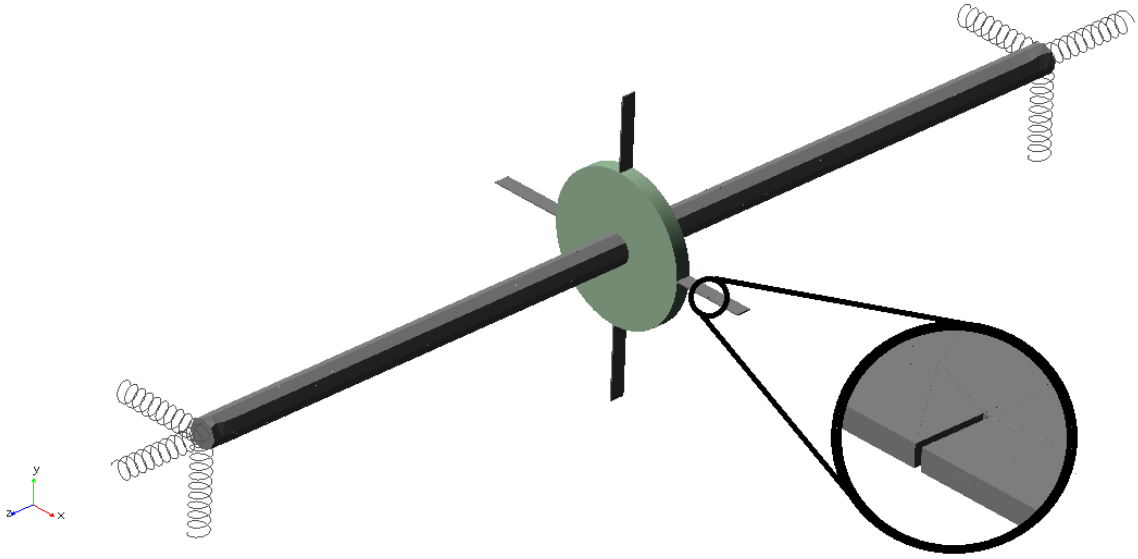


Figure 1. The configuration and crack modelling in the SDCB model.

It is assumed that the crack in the SDCB model is always open, and its edges have no contact. The shaft and the four blades of SDCB are assumed to be flexible. The effect of element numbers on the natural frequencies is checked during a convergence test, and the system is modeled using 4384 quadratic tetrahedral solid elements. Since the shaft and blade deformation relative to the disk is not comparable, the disk is modeled as a rigid body. Table 1 indicates the geometrical and material parameter values used in the SDCB system.

To ensure that the SDCB system is correctly modeled in MSC Adams, first, the present modeling procedure in this software and the Ansys workbench, as well as the assumed mode method (AMM), have been compared to a similar system without blades and crack. Table 2 shows that the system is correctly modeled in the MSC Adams software.

Table 1. Specifications of the SDCB system.

Parameter	Value	Parameter	Value
Shaft diameter [m]	0.060	Rotational speed [rad/s]	314
Shaft length [m]	2	Number of the blades	4
Disk diameter [m]	0.300	Blade thickness [m]	0.003
Disk thickness [m]	0.030	Blade length [m]	0.150
Young's modulus [GPa]	196	Blade width [m]	0.030
Poisson's ratio	0.22	Blade angle set [degree]	0
Density [Kg/m^3]	8180	Stiffness of the springs [MN/m]	10

Table 2. Comparison of the natural frequencies obtained from AMM, Ansys, and Adams software.

Natural Frequencies (Hz)	AMM	Ansys (Relative error)	MSC Adams (Relative error)
1st mode of the shaft bending	21.98	22.32 (1.55%)	21.93 (0.23%)
2nd mode of the shaft bending	113.01	112.41 (0.53%)	112.24 (0.68%)

Then, the SDCB system is modeled by adding the blades with geometrical properties mentioned in Table 1, and the crack is considered in one of the blades. The crack locations are assumed to be 20, 75, and 130 mm (from the blade root), and the crack depth is considered 5, 10, and 15 mm in each location. The crack with a smaller crack depth than 5 mm is not considered due to the low accuracy of the results. The signals were captured from two different locations with a sampling rate of 2048 Hz and 2 seconds time duration. The first location is a point near the blade tip, and the second is near

one of the bearings. The displacement and acceleration signals in the x and y axis are used in this study for the blade tip and one of the bearings, respectively. The two crack detection methods are applied to both signals to obtain the best method and location for capturing the signals. The following section describes this paper's two crack detection models in detail.

3. Crack detection classifiers

In this section, the automatic and manual feature selection procedures are described. The manual feature selection approach is described in sub-section 3.1, in which the SVM, Decision tree, and k-NN methods build an ensemble classifier. Then, the presented automatic MLP-based feature selector and classifier are defined in section 3.2.

Detection of the cracks located in the blades is not an easy task. In this paper, two points are considered to capture the signals. The first way is to install the accelerometer on one of the bearings. The second way is to use methods that measure the blade tip displacement. There are two approaches for this purpose. The first approach uses strain gauges, and the second is blade tip timing (BTT) measurement methods that have become very common in recent years [13,14].

Anyway, this measurement is done for ten different conditions and in x and y directions. These conditions include a healthy and cracked blade, where cracks are considered in three different locations and depths. So, 20 signals for 2 seconds and with a 2048 Hz sampling rate can be measured.

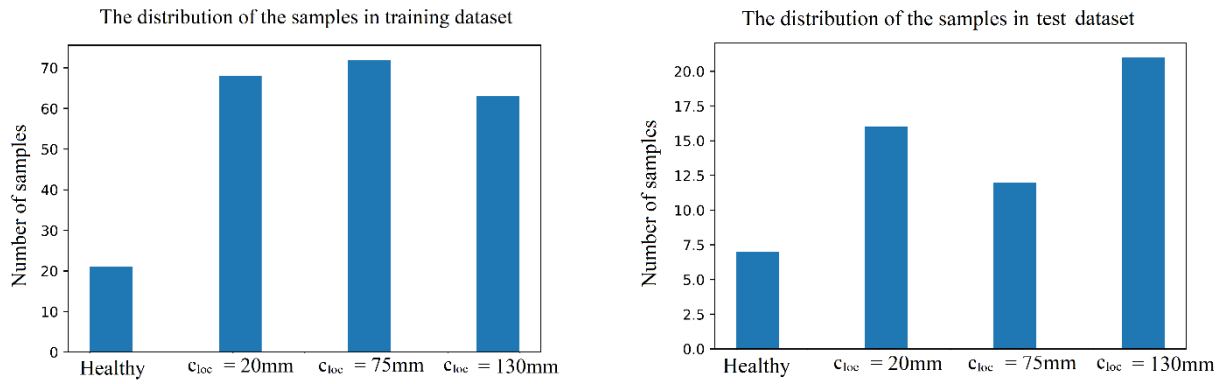
On the other hand, the 20 samples (signals) are not enough to train the model. Also, the signal is stationary since no damping is considered. In this situation, the data segmentation technique produces more samples from the original ones. So, in this study, the segmentation technique is used by sweeping a 512-point-length window with a step length of 256 points and generating new samples. Finally, the dataset size is increased from 20 to 280 samples. For the model training procedure, the 280 samples were split into 224 training samples (80% of the dataset) and 56 test samples (20% of the dataset).

Having 280 samples of length 512 (or 512 features) is not desirable. Many techniques are often used to reduce the amount of data for classification [13] and are called dimensionality reduction techniques. The fourteen features used instead of the 512-point signals are mentioned in Table 3.

Table 3. The features used as the model input.

	Feature name		Feature name
1	Root mean square (RMS)	8	Variance
2	Kurtosis	9	Trim-mean
3	Skewness	10	Sum
4	Mean	11	Trap-z
5	Standard deviation	12	Mad
6	Peak value	13	Mode
7	Min	14	Moment

Figure 2-a and -b shows the distribution of training and test datasets, respectively. In this figure, c_{loc} indicates the crack location. Figure 2-a depicts that 70 of 224 samples belong to the crack location of 75mm, which means that the occurrence probability is 31.25%. Similarly, for the test dataset, the maximum number of samples belongs to the crack location of 130mm, and the occurrence probability is 37.5%, which means that any model with an accuracy of less than 37.5% performs worse than the equal chance condition between the classification classes.



(a) (b)

Figure 2. The distribution of the (a) train dataset, and (b) test dataset

3.1 Crack detection using manually selected features

This section proposes an ensemble classifier including SVM, Decision tree, and k-NN classifiers. The ensemble classifier guarantees the generality and robustness of the model. The inputs of the ensemble classifier are selected manually from the fourteen features mentioned in Table 3.

The feature selection procedure is the crucial difference between this classifier and the MLP-based classifier, which is introduced in the next section. Since the features are selected manually, each of the classifiers' performance is dependent on the selected features. Some features improve the performance of the ensemble model, and others either have no effect on the performance or provide the worse performance along with other features. The reason for this may be that features are used with equal importance. Some of them may be less important than others.

Because of the different well-known classifiers such as SVM, Decision tree, and k-NN, the manually selected features results in different accuracy. To prevent the poor accuracy of crack detection, one can decide to combine them. Therefore, a more robust and stable classifier is built.

3.2 Crack detection using MLP-based automatic feature selector

In this section, the fourteen features were entered into a deep multi-layer perceptron neural network classifier. Then, the weights of the network are automatically calculated to indicate each feature's importance to improve the accuracy of the crack detection. The weights of neural networks may become near zero for a specific feature, its means that feature has no effective information. Therefore, an important disadvantage of the ensemble classifier is not existed in the MLP-based classifier and it is expected to perform better.

One can consider the dataset as a matrix which rows and columns indicate the number of samples and the number of features. Using the segmentation technique, the size of the dataset matrix is changed from 20×4096 to 280×512. Next, the 14 features are calculated for each row, and the dataset size becomes 280×14. Finally, this dataset is split into 224 samples (the training dataset) and entered into the MLP-based classifier. The specification and techniques used in this classifier are indicated in Figure 3.

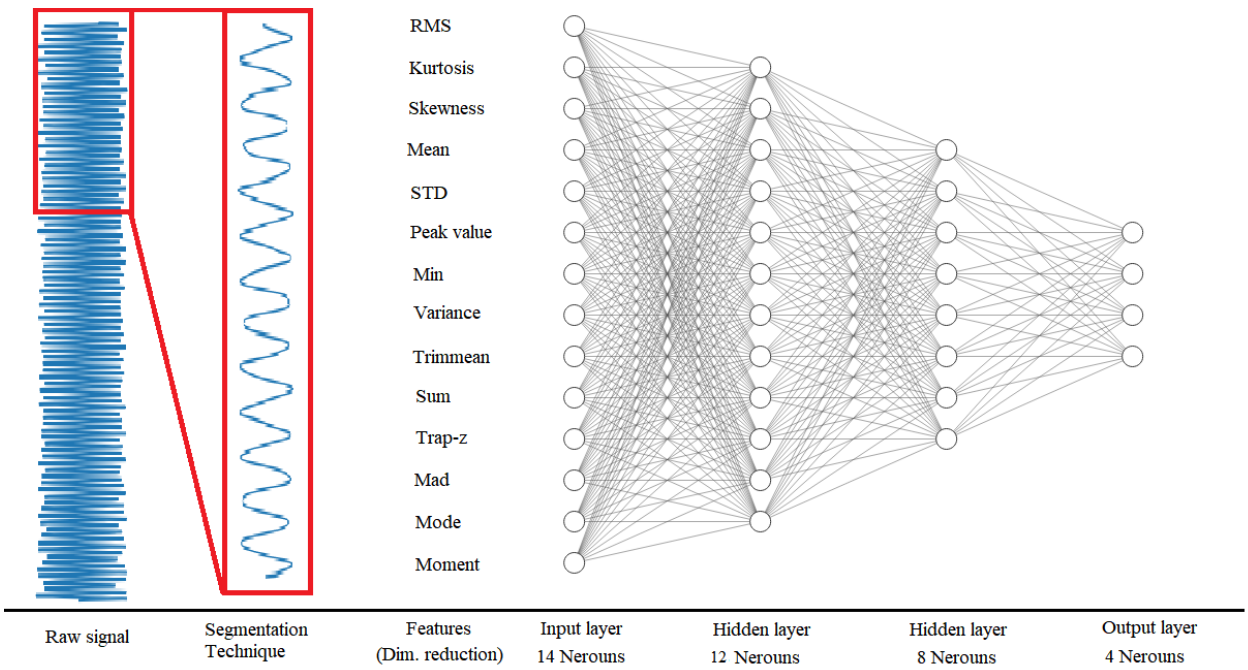


Figure 3. Schematic of the deep MLP-based classifier

In this network, the activation function of all layers except the last layer is the rectifier linear unit (ReLU) as a nonlinear activation function. Therefore, the network can capture the nonlinear property of the signals in deep neural network architecture due to the presence of the crack. The key benefit of employing the ReLU function over other activation functions is that it does not stimulate all neurons simultaneously. The activation function of the last layer is softmax. The softmax function is commonly utilized as the neural network's last activation function to normalize the network's output to a probability distribution over the expected output class. As a result of using softmax activation in the last layer, the network works as a logistic regression classifier. The MLP-based classifier is trained for the maximum number of iterations of 4000, the epoch size is 32, and the early stopping technique is used to avoid the overfitting of the model, so the model is not over-trained.

4. Results and discussion

The detection of crack location is the primary goal of this paper. The manual and automatic feature selection ability to detect the crack location are mentioned in sub-sections 4.1 and 4.2, respectively. Also, the classifier's performance using two datasets collected from the point near the cracked blade tip and a point near the bearing are compared.

4.1 Ensemble classifier with manual feature selection

After examining various feature sets, the best features were selected as the kurtosis, skewness, standard deviation, peak value, minimum, and moment. Table 4 shows the performance of the SVM, Decision tree, k-NN, and ensemble classifier using these six features for the displacement signal of the blade tip.

As it can be observed from Table 4, using the six features mentioned earlier, the Decision tree provides a more accurate result for this research, but it is not a general conclusion. The most crucial information someone can achieve from the comparison of the various classifier is to learn more about the dataset distribution. In this study, the k-NN classifier performs far better in classifying the blade's crack location than the SVM classifier, but its accuracy is less than the Decision tree classifier. The

SVM classifier uses hyperplanes to separate the classes. In this investigation, the hyperplanes of the sci-kit-learn python package are used. Therefore, it is evident that the dataset is not easily separable using each of the hyperplanes of the SVM classifier in this package. If one can introduce a custom better hyperplane, it should result in better test accuracy. However, before introducing a new hyper-plane, another important note remains. The input of such a classifier is manually selected, and this may be the main reason for the relatively low accuracy of the SVM method. In this study, SVM uses linear hyperplanes, so the dataset (or selected features) is not linearly separable.

Another important conclusion is that the use of the classifiers is dependent on the selected features, and it can be the reason for the low performance of a powerful classifier (such as SVM in this investigation). Therefore, the combination or ensemble of the various classifiers can have more robust and stable results.

Table 4. The accuracy of the classifiers

Name of the classifier	Train accuracy	Test accuracy
SVM	32.14%	21.43%
Decision tree	97.32%	89.29%
k-NN	100%	83.93%
Ensemble classifier (Includes all of the above classifiers)	100%	85.71%

The confusion matrix, a matrix that shows the accuracy of the prediction of each class, for the ensemble classifier is illustrates in Figure 4 for displacement signal of the blade tip. The accuracy of the ensemble classifier is calculated according to the diagonal correct prediction accuracy (in green) and the number of samples of each class in the test dataset.

		Predicted labels			
		Healthy	c_{loc} = 20mm	c_{loc} = 75mm	c_{loc} = 130mm
Actual labels	Healthy	57.14	14.29	0.00	28.57
	c_{loc} = 20mm	6.25	75.00	18.75	0.00
	c_{loc} = 75mm	0.00	0.00	100.00	0.00
	c_{loc} = 130mm	0.00	4.76	0.00	95.24

Figure 4. The confusion matrix of the crack detection from displacement of cracked blade tip

The dataset from the acceleration signals of a point near one of the bearings is considered to identify a better point for data acquisition purposes, and the same crack detection method is done on that dataset. The accuracy of the classifiers is shown in Table 5, and the confusion matrix is indicated in Figure 5.

Table 5. The accuracy of the classifiers

Name of the classifier	Train accuracy	Test accuracy
SVM	42.41%	42.86%
Decision tree	82.59%	42.86%
k-NN	100%	57.14%
Ensemble classifier (Includes all of the above classifiers)	90.62%	48.21%

According to Table 5 and Figure 5, it is evident that the performance of ensemble classifier using dataset collected from the tip of the blade is far better than the same classifier using acceleration of one of the bearings.

		Predicted labels			
		Healthy	c_{loc} = 20mm	c_{loc} = 75mm	c_{loc} = 130mm
Actual labels	Healthy	100.00	0.00	0.00	0.00
	c_{loc} = 20mm	0.00	68.75	18.75	12.50
	c_{loc} = 75mm	0.00	33.33	25.00	41.67
	c_{loc} = 130mm	0.00	52.38	19.05	28.57

Figure 5. The confusion matrix of the crack detection from acceleration of one of the bearings

4.2 MLP-based classifier with automatic feature selection model

Despite the ensemble classifier, the deep MLP-based classifier uses all of the fourteen features mentioned in Table 3 and automatically determine the importance of each of them using the network weights. The train and test loss value reduction during the training procedure is illustrated in Figure 6 for one of the training procedures.

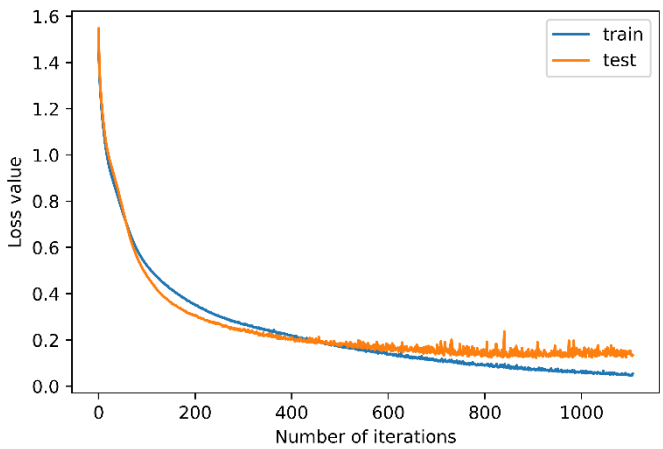


Figure 6. Train and test loss tracking for displacement signals collected from the tip of the cracked blade

It is evident that the loss value of the test dataset is consistent after about 800 iterations and tends to increase, but the loss value of the training dataset is reducing. So, the early stopping technique eliminates the training procedure to prevent the overfitting of the model.

According to the confusion matrix (Figure 7), the model predicts the healthy blade and cracks located at 130 mm of the blade root without an error. However, when the crack location is 20 or 75mm from the blade root, there is a 6.25% and 8.33% probability of misunderstanding of crack location. The blade length is 150mm according to Table 1 and can be considered a cantilever beam.

In such a beam, a crack can be the reason for locally reducing beam stiffness. The more the crack is close to the free edge of the beam, the more the effect of the crack is sensitive (due to its high deflection). Therefore, it is reasonable that detecting the crack close to the free edge of the beam is easier than the crack in the middle of the beam. As a result, it can be seen in Figure 7 that the accuracy of a healthy blade and a cracked blade with a crack location of 130mm is 100%.

		Predicted labels			
		Healthy	c_{loc} = 20mm	c_{loc} = 75mm	c_{loc} = 130mm
Actual labels	Healthy	100.00	0.00	0.00	0.00
	c_{loc} = 20mm	0.00	93.75	6.25	0.00
	c_{loc} = 75mm	0.00	8.33	91.67	0.00
	c_{loc} = 130mm	0.00	0.00	0.00	100.00

Figure 7. The confusion matrix of the crack detection from acceleration of near cracked blade tip point

The loss value for train and test datasets during the training procedure (Figure 8) for the bearing acceleration signals shows that the two curves diverge after about 125 iterations which is very low compared to Figure 6.

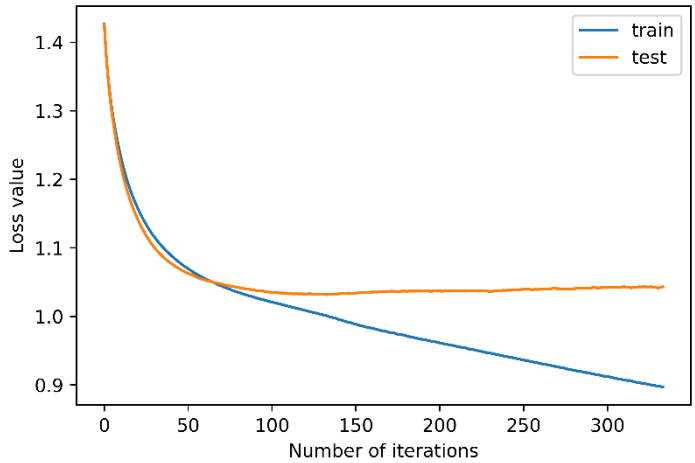


Figure 8. Train and test loss tracking for acceleration signals collected from one of the bearings

The model performance for the acceleration signals collected from one of the bearings is shown in the confusion matrix of Figure 9. In this figure, the healthy state is detected correctly, but the crack location is not detected. Although, the measuring of the bearing acceleration signals is a proper way to find the crack presence in the system (100% accuracy), but it is not a proper way to realize the location and other properties of the crack. Therefore, it is necessary to use signals from the blade tip.

The training and testing of the model are repeated ten times for each dataset. Then the average performance of the model is reported in Table 6.

		Predicted labels			
		Healthy	c_{loc} = 20mm	c_{loc} = 75mm	cloc = 130mm
Actual labels	Healthy	100.00	0.00	0.00	0.00
	c_{loc} = 20mm	0.00	31.25	25.00	43.75
	c_{loc} = 75mm	0.00	33.33	33.33	33.33
	c_{loc} = 130mm	0.00	33.33	19.05	47.62

Figure 9. The confusion matrix of the crack detection from acceleration of near bearing point

Table 6. The performance of the MLP-based classifier for both datasets

Dataset	Train accuracy	Test accuracy
The acceleration of a point near one of the bearings	61.96%	43.03%
The displacement of the cracked blade tip	99.42%	97.14%

According to Table 6, the crack detection results show poor accuracy when the acceleration signals are collected from a point near one of the bearings. On the contrary, the crack detection method based on a deep MLP network performs very well, and the average accuracy reaches 97.14%.

5. Conclusion

Most studies in fault diagnosis have focused almost on the bearings, gearbox, and rotor faults, and there are inadequate papers on automatic crack detection using neural networks. Detection of the crack in a system has always been a challenge. Therefore, this paper investigated the benefits of using automatic feature selection in a deep MLP-based classifier to detect crack location in a basic shaft-disk-cracked blade system. Three crack locations with three different depths were chosen. This study used two kinds of signals to show the importance of the sensor's location. One of them is the acceleration signals of a point near one of the bearings, and another is the displacement of the cracked blade tip. The latter signals can be captured using BTT methods or strain gauges, which is practically possible. The results showed that the average accuracy of the MLP-based classifier reached 97.14% for the blade tip signals, but for acceleration signals of the bearing, the accuracy was 43.03%. For comparison, an ensemble classifier using SVM, Decision tree, and k-NN classifiers was introduced, and its accuracy reached 85.71% for blade tip signals and 48.21% for the acceleration of the bearing. From the results, it can be concluded that the accuracy was extensively dependent on the type of classifiers, the feature extraction and selection procedure, and the location of the sensors. One of the most significant achievements of this investigation is that installing sensors that can capture the blade tip displacement signals is more beneficial in detecting the location of the crack. Another important conclusion is that the automatic selection of features with different contributions to the classification can be a better approach to identifying the crack location. As a result, this paper suggests using the proposed MLP-based classifier. There is also some limitation to using the presented MLP-based classifier. The model cannot consider the dataset containing a crack depth of less than 5 mm, leading to lower accuracy results.

REFERENCES

[1] T. Zhang, J. Chen, F. Li, K. Zhang, H. Lv, S. He, and E. Xu, "Intelligent fault diagnosis of machines with small & imbalanced data: A state-of-the-art review and possible extensions," *ISA Trans.* **119**, 152–171 (2022).

[2] H. Alavi, A. Ohadi, and S. T. Niaki, "A novel targeted method of informative frequency band selection based on lagged information for diagnosis of gearbox single and compound faults," *Mech. Syst. Signal Process.* **170**, 108828 (2022).

[3] B. Li, B. Tang, L. Deng, and J. Wei, "Joint attention feature transfer network for gearbox fault diagnosis with imbalanced data," *Mech. Syst. Signal Process.* **176**, 109146 (2022).

[4] M. Gohari and A. M. Eydi, "Modelling of shaft unbalance: Modelling a multi discs rotor using K-Nearest Neighbor and Decision Tree Algorithms," *Measurement* **151**, 107253 (2020).

[5] P. Shinde and R. G. Desavale, "Application of dimension analysis and soft competitive tool to predict compound faults present in rotor-bearing systems," *Measurement*, 110984 (2022).

[6] Y. Liu, J. T. Li, K. P. Feng, Y. L. Zhao, X. X. Yan, and H. Ma, "A novel fault diagnosis method for rotor rub-impact based on nonlinear output frequency response functions and stochastic resonance," *J. Sound Vib.* **481**, 115421 (2020).

[7] Q. Wang, W. Wu, F. Zhang, and X. Wang, "Early rub-impact fault detection of rotor systems via deterministic learning," *Control Eng. Pract.* **124**, 105190 (2022).

[8] S. Fatima, A. R. Mohanty, and V. A. Naikan, "A misalignment detection methodology by measuring rate of temperature rise of shaft coupling using thermal imaging," *Proc. Inst. Mech. Eng. Part O J. Risk Reliab.* **229**, 209–219 (2015).

[9] T. Zhang, F. Xu, and M. Jia, "A centrifugal fan blade damage identification method based on the multi-level fusion of vibro-acoustic signals and CNN," *Measurement* **199**, 111475 (2022).

[10] T. Ma, F. Xu, J. Hu, D. Song, and S. Cao, "Double Gaussian potential stochastic resonance method and its application in centrifugal fan blade crack detection," *Chin. J. Phys.* **74**, 279–295 (2021).

[11] F. Mevissen and M. Meo, "Ultrasonically stimulated thermography for crack detection of turbine blades," *Infrared Phys. Technol.* **122**, 104061 (2022).

[12] Y. Zhang, F. Avallone, and S. Watson, "Wind turbine blade trailing edge crack detection based on airfoil aerodynamic noise: An experimental study," *Appl. Acoust.* **191**, 108668 (2022).

[13] S. Chen, Y. Yang, H. Hu, F. Guan, G. Shen, Z. Bian, and H. Guo, "Interpolation method for wideband signal reconstruction based on blade tip timing measurement," *Measurement* **176**, 109168 (2021).

[14] P. Beauseroy and R. Lengellé, "Nonintrusive turbomachine blade vibration measurement system," *Mech. Syst. Signal Process.* **21**, 1717–1738 (2007).